

Rapid Data Backup and Restore Using NFS on IBM ProtecTIER TS7620 Deduplication Appliance Express

IBM Redbooks Solution Guide

This IBM® Redbooks® Solution Guide describes how data backup and restore processes can be speeded up by using the IBM TS7620 ProtecTIER® Deduplication Appliance Express and the newly available Network Attached Storage (NAS) support for Linux- and UNIX-based client backup environments. The File System Interface (FSI) with NFS support helps facilitate rapid data restoration, and the IBM HyperFactor® algorithm helps maximize disk usage by eliminating data duplicates inline from the incoming backup data streams. The TS7620 Appliance Express comes pre bundled with internal disk storage included in a customer-installable, ready-to-run configuration. The ProtecTIER system can be integrated into an existing backup solution, and provides deduplication when saving files with backup solutions such as IBM Tivoli® Storage Manager, Symantec Netbackup, EMC Networker, and IBM i Backup and Recovery and Media Services (BRMS).

Figure 1 shows a solution that uses Tivoli Storage Manager to back up to and restore data from the self-contained disk storage in the TS7620 Deduplication Appliance Express. The TS7620 is configured for FSI-NFS, which emulates UNIX file system behavior and presents a virtualized hierarchy of file systems directories and files to UNIX-based clients using NFS protocol. The ProtecTIER Manager graphical user interface (GUI) is used to monitor the TS7620 Deduplication Express. With ProtecTIER Version 3.3, this GUI can be used to upgrade the ProtecTIER system to new code levels.

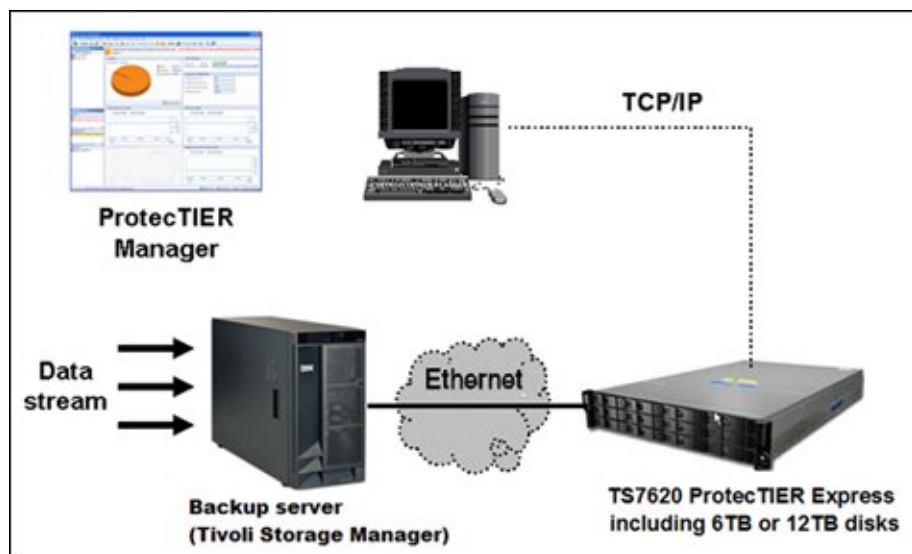


Figure 1. Using an IBM ProtecTIER TS7620 Deduplication Appliance Express with FSI support

Did you know?

- FSI enables backup servers to run backup applications to the ProtecTIER appliance over IP networks, without requiring Fibre Channel connectivity. Using NFS to support your virtual server environment over IP networks provides advantages over using Fibre Channel in terms of cost and complexity.
- FSI support enables existing tools and capabilities to integrate a ProtecTIER NAS-based backup target into the existing environment.
- No knowledge of tape and drive interfaces is required. FSI support provides a file-based interface.
- CIFS and NFS can coexist in the same repository.

Business value

ProtecTIER deduplication can be used to perform faster, more frequent backups and quickly restore backed up data to maintain business continuity.

Using data deduplication, you can reduce the amount of space that is required to store data on disk. With deduplication, repeated instances of identical data are identified and stored in a single instance. This process saves storage capacity and bandwidth. Data deduplication can provide greater data reduction than previous technologies, such as Lempel-Ziv (LZ) compression and differencing, which is used for differential backups.

Data deduplication is performed while the data is being backed up to the ProtecTIER server (*inline* deduplication), in contrast to after the data is written to the server (*post processing*). The advantage of inline data deduplication is that the data is processed only once; no additional processing is needed after the backup window. Inline data deduplication requires less disk storage because the native data is not stored before data deduplication.

Depending on your configuration, this solution can provide some or all of the following benefits:

- Greater productivity because more frequent backups can be performed with the existing physical disk space
- Reduced energy consumption, which results from reducing the amount of disk in operation
- A smaller footprint, which reduces the amount of physical space that is required for the storage solution
- Fast restoration, which can often be more critical to business continuity than fast backup
- ProtecTier deduplication can reduce data storage by as much as 96%
- Disaster resilience using network replication; the bandwidth requirement is reduced because only the deduplicated data must be replicated
- Improved performance for frequent full backups because ProtecTIER servers have the best deduplication, the highest backup speed, and the highest restore speed for multiple full backups of the same objects

This solution enables you to take advantage of HyperFactor® technology, which uses a pattern algorithm that can reduce the amount of space required for storage of the backup environment by a factor of up to 25 times, based on evidence from existing implementations. The effect and result of HyperFactor processing is a *factoring ratio*. In simple terms, the factoring ratio is the ratio of nominal data (as a sum of all user data backup streams) to the occupied physical storage in the ProtecTIER repository. For example, a 10:1 ratio means that 10 times more nominal data is being managed than the physical space required to store it. Also, the 10:1 ratio means a 90% storage savings. The effectiveness of data deduplication depends on many variables, including the rate of data change, the number of backups, and the data retention period.

Figure 2 shows the relationship between HyperFactor processing and the possible storage savings.

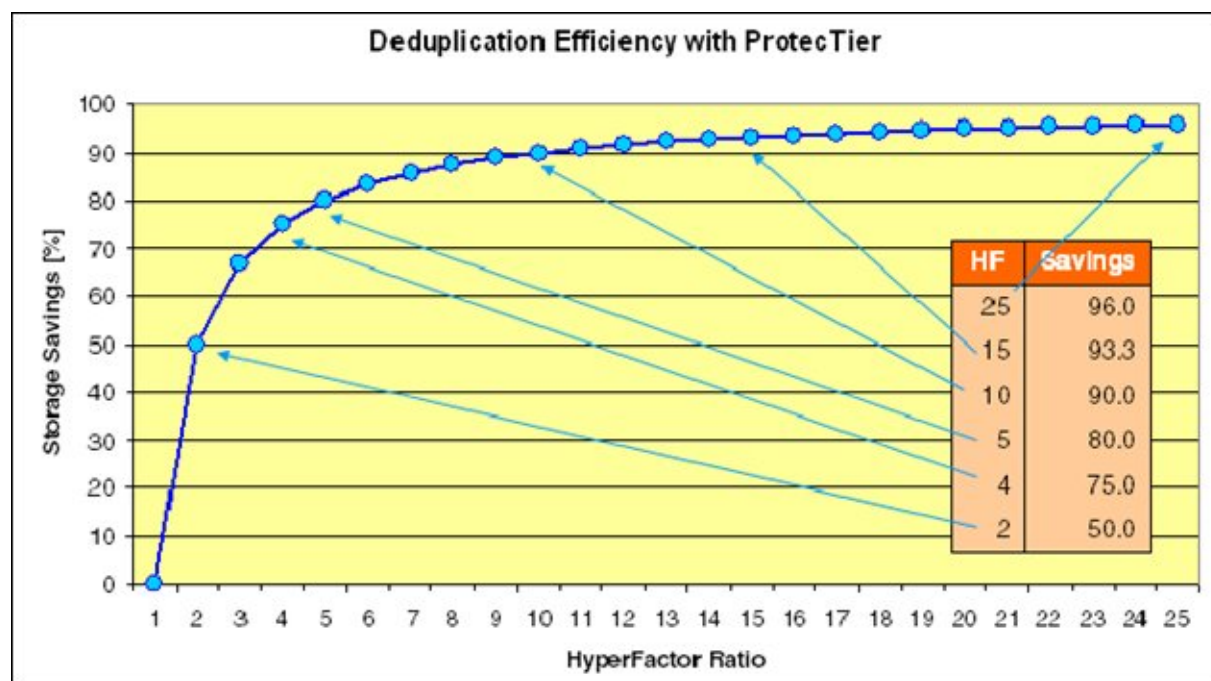


Figure 2. Storage savings and deduplication ratio

Solution overview

A ProtecTIER system is configured to use FSI for deduplication and backup. The ProtecTIER system presents itself as a network-attached storage (NAS) backup target that is capable of using the HyperFactor algorithm and ProtecTIER native replication bandwidth reduction techniques for storing and replicating deduplicated data to make backup and WAN vaulting more cost effective. FSI enables backup servers that are running backup applications to connect over IP networks to ProtecTIER, without requiring Fibre Channel connectivity.

ProtecTIER for FSI supports two file systems and network protocols: Network File System (NFS) and Common Internet File System (CIFS).

- NFS is used primarily for Linux- or UNIX-based operating systems and provides remote access. Files on one computer can be viewed or edited from a remote computer.
- CIFS file sharing is Windows-based. The CIFS protocol allows join sharing of multiple devices, such as printers, files, and serial ports, between multiple users and administrators. The CIFS protocol supports remote mounts over TCP/IP that is using the Samba Server Message Block (SMB) protocol.

The ProtecTIER Manager 3.3 software includes FSI support for Windows-based servers as CIFS clients and UNIX-based servers as NFS exports. ProtecTIER emulates Windows file system behavior and presents a virtualized hierarchy of file systems, directories, and files to Windows CIFS clients. These Windows CIFS clients are able to perform normal Windows file system operations on the emulated file system content.

The ProtecTIER Manager software includes interfaces for:

- Managing network connections for the ProtecTIER backup server
- Managing FSI network connections
- Managing users and file systems

- Creating and managing NFS and FSI shares
- Simulating disaster recovery

Figure 3 shows the interface.

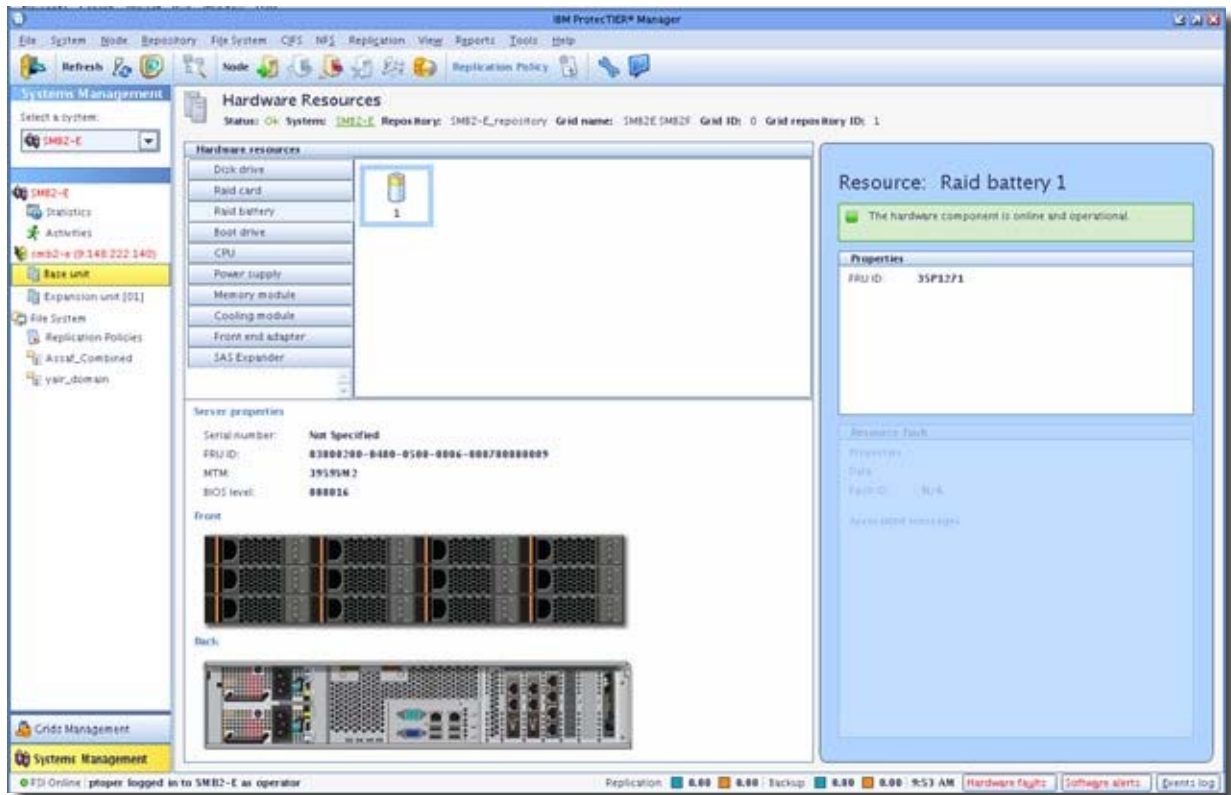


Figure 3. ProtecTIER Manager interface

Solution architecture

This solution uses a TS7620 Appliance as a deduplication server for FSI backup targets. The ProtecTIER TS7620 Deduplication Appliance runs Tivoli® Storage Manager as a backup application solution. If FSI-CIFS is installed, FSI-NFS can be installed on the same system. Any directory that is written or created as CIFS can be exported as an NFS mount point, and vice versa.

The TS7620 Appliance system is a stand-alone configuration (with one node), and up to two expansion units. It uses one server to transfer data from the backup server to the TS7620 internal storage.

The ProtecTIER File System Interface (FSI) presents ProtecTIER as a network-attached storage (NAS) backup target capable of using the HyperFactor algorithm and ProtecTIER native replication bandwidth reduction techniques for storing and replicating deduplicated data. The ProtecTIER FSI interface is intended to be used for backup and restore of data sets using a backup application.

Usage scenarios

Scenario 1

This use case consists of a very simple environment running on a Linux or UNIX system. It has a Tivoli Storage Manager Server attached to three Tivoli Storage Clients, which perform backups and restores to the IBM Storage TS7620 Deduplication Appliance Express (TS7620 Appliance) configured as an FSI-NFS environment. There are no physical cartridges.

Figure 4 shows the solution environment.

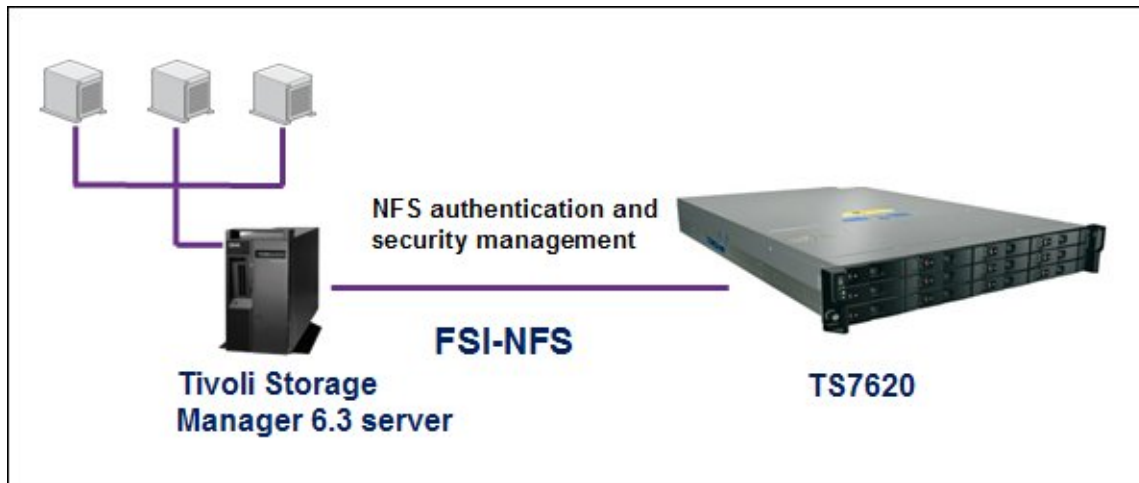


Figure 4. Simple environment with a Tivoli Storage Manager Server attached to three Tivoli Storage Clients

Scenario 2

This use case illustrates an environment with two TS7620 systems. One system is used as a primary site and the second system is used as a disaster recovery site. The system is configured as a Linux or UNIX system with a Tivoli Storage Manager Server attached to three Tivoli Storage Clients with the following attributes:

- A GMU customer has a primary site replicating 30 miles away.
- The customer is using the primary TSM while in data recovery testing.
- In a real disaster, a stand-by server is used.
- There are no physical cartridges. TSM backs up to and restores from an NFS file share.

Figure 5 shows this environment with two TS7620 servers: GMU1 and GMU2.

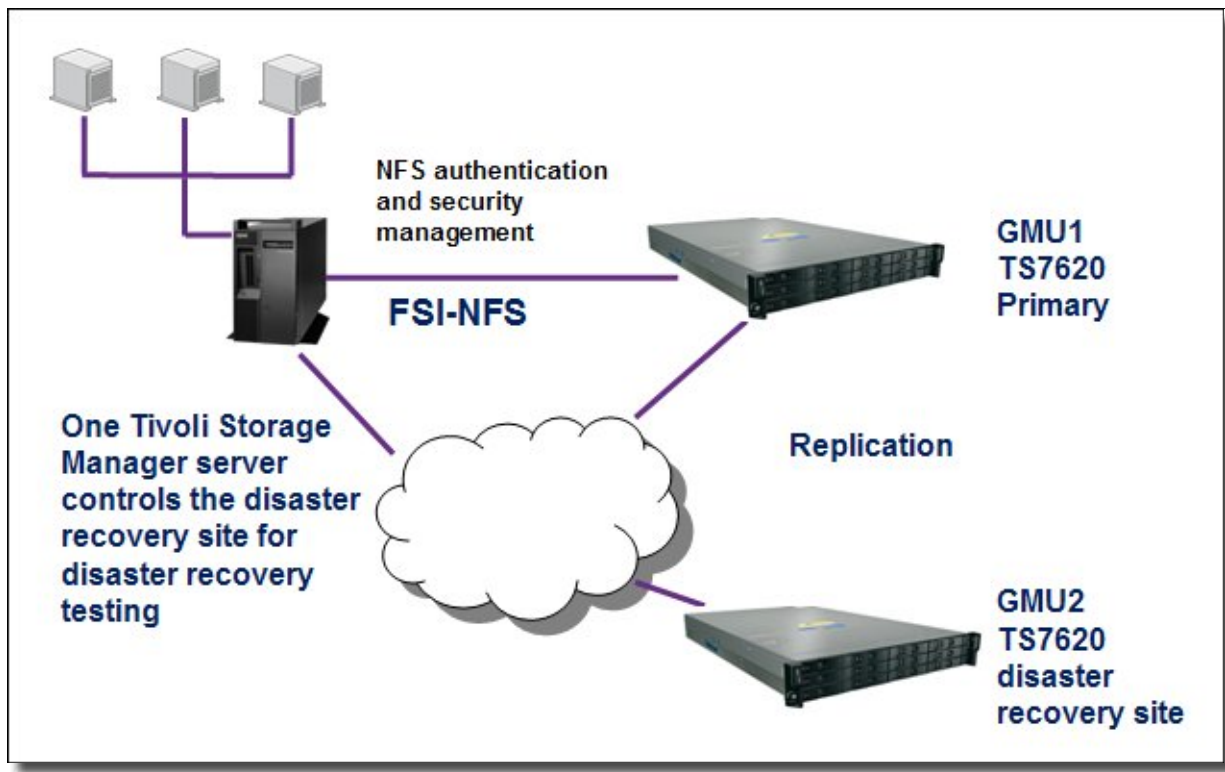


Figure 5. Example environment with two Tivoli TS7620 servers

Integration

The solution requires an IBM ProtecTIER TS7620 appliance, the IBM System Storage® ProtecTIER Entry Edition software Version 3.3, and IBM Tivoli Storage Manager Version 6.3. The ProtecTIER TS7620 appliance can be integrated into an existing TCP/IP network.

Supported platforms

This solution requires an IBM ProtecTIER TS7620 appliance and ProtecTIER Manager software.

The ProtecTIER Manager (PT Manager) application is a graphical user interface (GUI) that you use to view, configure, manage, upgrade, and troubleshoot the operation of the IBM TS7600 ProtecTIER family of products. The workstation for the installation of the PT Manager is not included in the ProtecTIER system.

The IBM ProtecTIER Manager software can be installed on x86 hardware that is running one of the following 32- or 64-bit operating systems:

- Windows operating system (Windows 2003, Windows XP, or Windows 7)
- RedHat Linux (Enterprise 4 or higher)

Ordering information

Table 1 shows the program numbers for the IBM products that are used in this solution.

Table 1. Ordering part numbers and feature codes

Program name	Description	Program number/version
ProtectTIER V3.3 Entry Edition preload	This feature must be selected to instruct IBM manufacturing to install the ProtectTIER Entry Edition software on the TS7620 server. The software (5639-PTC) must be ordered separately.	9318 Version 3.3
IBM System Storage TS7620 ProtecTIER Deduplication Appliance	This feature must be selected to instruct IBM manufacturing to install the ProtectTIER Entry Edition software on the TS7620 server. The software (5639-PTC) must be ordered separately.	3959-SM2 N/A
IBM System Storage TS7620 ProtecTIER Deduplication Appliance	Optional expansion unit. For this solution, this expansion can be used only with the 3959-SM2 appliance	3959-EXP N/A
IBM Tivoli Storage Manager	Tivoli Storage Manager is a centralized policy-based data backup and recovery software. The software enables a user to back up, restore, archive, and retrieve data from a hierarchy of data storage areas. The storage areas, known as pools, can be a hierarchy of disk, optical, and tape-based media.	5680-E01 Version 6.3

A TS7620 base unit with 6 TB capacity can be upgraded to a 12 TB model with a software upgrade. No additional hardware is required. For additional capacity upgrades, up to two 3959-EXP expansion units can be added.

Related information

For more information, see the following documents:

- *IBM System Storage TS7600 with ProtecTIER Version 3.3*, SG24-7968
- *IBM ProtecTIER implementation and Best Practices Guide*, SG24-8025
- *Implementing IBM Storage Data Deduplication Solutions*, SG24-7888
- *IBM System Storage TS7620 ProtecTIER Deduplication Appliance Introduction and Planning Guide*, GA32-0913
- *IBM System Storage TS7600 - ProtecTIER v.3.3 User's Guide for FSI Systems*, GA32-2235
- *IBM System Storage TS7600 with ProtecTIER User's Guide for Enterprise Edition and Appliance Edition*, GC53-1156
- *IBM System Storage TS7620 ProtecTIER Deduplication Appliance Express Introduction and Planning Guide for VTL, OpenStorage, and FSI systems*, GA32-0913
- *IBM System Storage TS7620 ProtecTIER Deduplication Appliance Express Installation and Setup Guide for Installation and Setup Guide for VTL, OpenStorage, and FSI systems*, GA32-0914
- *IBM System Storage TS7610 and TS7620 ProtecTIER Deduplication Appliance Express V3.3, Maintenance Guide*, GA32-223.
- *IBM System Storage TS7610 or TS7620 ProtecTIER v3.3, Deduplication Appliance Express V3.3, Software Upgrade Guide*, SC27-3641

- *IBM System Storage TS7610 and TS7620 ProtecTIER Deduplication Appliance Express V3.3, Service Guide*, GA32-0915
- IBM TS7620 ProtecTIER Deduplication Appliance Express
<http://www.ibm.com/systems/storage/tape/ts7620>
- IBM Offering Information page (announcement letters and sales manuals):
http://www.ibm.com/common/ssi/index.wss?request_locale=en

On this page, enter *IBM System Storage TS7620*, select the information type, and then click **Search**. On the next page, narrow your search results by geography and language.

Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service. IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:

IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785 U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you. This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk. IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you. Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products. This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurement may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs.

© Copyright International Business Machines Corporation 2013. All rights reserved.

Note to U.S. Government Users Restricted Rights -- Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

This document was created or updated on April 3, 2013.

Send us your comments in one of the following ways:

- Use the online **Contact us** review form found at:
ibm.com/redbooks
- Send your comments in an e-mail to:
redbook@us.ibm.com
- Mail your comments to:
IBM Corporation, International Technical Support Organization
Dept. HYTD Mail Station P099
2455 South Road
Poughkeepsie, NY 12601-5400 U.S.A.

This document is available online at <http://www.ibm.com/redbooks/abstracts/tips0990.html> .

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. These and other IBM trademarked terms are marked on their first occurrence in this information with the appropriate symbol (® or ™), indicating US registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at <http://www.ibm.com/legal/copytrade.shtml>. The following terms are trademarks of the International Business Machines Corporation in the United States, other countries, or both:

HyperFactor®
IBM®
ProtectTIER®
Redbooks (logo)®
System Storage®
Tivoli®

The following terms are trademarks of other companies:

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

Windows, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Other company, product, or service names may be trademarks or service marks of others.